

Impedance adaptation for optimal robot–environment interaction

Article (Accepted Version)

Ge, Shuzhi Sam, Li, Yanan and Wang, Chen (2013) Impedance adaptation for optimal robot–environment interaction. *International Journal of Control*, 87 (2). pp. 249-263. ISSN 0020-7179

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/72085/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

RESEARCH ARTICLE

Impedance Adaptation for Optimal Robot-Environment Interaction

Shuzhi Sam Ge^{a,c*}, Yanan Li^{a,b} and Chen Wang^a

^a*Social Robotics Laboratory, Interactive Digital Media Institute, and Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119077*

^b*NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore 119613*

^c*Robotics Institute, and School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China*

(Received 00 Month 201x; final version received 00 Month 201x)

In this paper, impedance adaptation is investigated for robots interacting with unknown environments. Impedance control is employed for the physical interaction between robots and environments, subject to unknown and uncertain environments dynamics. The unknown environments are described as linear systems with unknown dynamics, based on which the desired impedance model is obtained. A cost function that measures the tracking error and interaction force is defined, and the critical impedance parameters are found to minimize it. Without requiring the information of the environments dynamics, the proposed impedance adaptation is feasible in a large number of applications where robots physically interact with unknown environments. The validity of the proposed method is verified through simulation studies.

Keywords: robot-environment interaction; impedance adaptation; optimal control

1 Introduction

In social applications such as elderly care, health care and human-robot collaboration, environments are typically unknown to robots and there exist uncertainties due to many factors. Therefore, control of interaction between robots and environments is essential and there has been much effort made on this topic. In the literature of interaction control, two approaches are widely studied, i.e., hybrid position/force control Craig and Raibert (1979), Huang et al. (2006), Li et al. (2007) and impedance control Hogan (1985). Compared to hybrid position/force control, impedance control is more feasible in the sense that it does not require the decomposition of two directions. Besides, impedance control is preferred for its better robustness Colgate and Hogan (1988).

Under impedance control, robots are governed to be compliant to the interaction force exerted by environments and thus the safety of both robots and environments can be guaranteed. Specifically, imposing a passive impedance model to robots will guarantee the interaction stability if environments are also passive Colgate and Hogan (1988). In the early research works of impedance control, a desired passive impedance model is usually prescribed and then the effort is focused on handling the uncertainties in robots dynamics. These works include adaptive impedance control such as Colbaugh et al. (1991), Lu and Meng (1991) and learning impedance control such as Cheah and Wang (1998), Wang and Cheah (1998), Li et al. (2012). However, in many situations, to impose a passive impedance model to the robot is too conservative, and the environments dynamics can be taken into consideration to obtain desired impedance model Buerger and Hogan (2007). Besides, a fixed prescribed impedance model does not

*Corresponding author. Tel: (+65) 6516 6821, Fax: (+65) 6779 1103, E-mail: samge@nus.edu.sg

suffice in many applications. For example, variable impedance control is necessary in human-robot collaboration Tsumugiwa et al. (2001, 2002) and explosive movement Braun et al. (2012a,b). Although the methods discussed in Tsumugiwa et al. (2001, 2002) provide a better control performance in the sense of more efficient human-robot collaboration, the resulted impedance parameters (mass, damping and stiffness) are obtained in a heuristical way and cannot be easily extended to other applications. To cope with this problem, iterative learning has been studied to obtain impedance parameters subject to unknown environments in an analytic way. It has been generally acknowledged that such an ability to improve performance by repeating a task is an important control strategy of the human being De Roover et al. (2000), Xu et al. (2000). Pioneered by Arimoto et al. (1984a,b), iterative learning control has been widely investigated for robot control. In Cohen and Flash (1991), associative search network is adopted for the impedance learning and the resulted impedance control is applied to a wall-following task. In Yang and Asada (1996), an internal model based impedance learning method is developed and used in a high-speed insertion application. In Tsuji and Morasso (1996), neural networks are employed to update both the impedance parameters and rest position iteratively. Compared to iterative impedance learning discussed above, impedance adaptation is more interesting yet it is more challenging. It is interesting because it does not require the robot to repeat operations to learn the desired impedance parameters. This is important because to make the robot repeat operations may cause inconvenience in many situations. It is challenging because to develop an adaptive scheme usually requires that a certain variable is invariant but this is difficult to satisfy in the case of dynamically changing environment. There has been research effort on impedance adaptation in the literature, although it is less compared to that on impedance learning. In Uemura and Kawamura (2009), stiffness is updated to minimize the actuator torque by taking resonance into consideration. In Stanisic and Fernandez (2012), the switching strategies of impedance parameters are discussed in order to dissipate the system energy and realize a “soft” interaction.

In the development of impedance learning and adaptation, optimization plays an important role because the control objective of impedance control includes both the force regulation and trajectory tracking and usually it is the compromise of these two objectives. In the case of impedance adaptation, a cost function or a reward function is defined to describe the interaction performance, and impedance parameters are expected to be adjusted to minimize the cost function or maximize the reward function. In Johansson and Spong (1994), the well-known linear quadratic regulator (LQR) is utilized to determine the desired impedance parameters with the environment dynamics known *a priori*. In Matinfar and Hashtrudi-Zaad (2005), impedance parameters are adjusted as the online solutions of the defined LQR problem, instead of fixing the impedance parameters obtained based on LQR as in Johansson and Spong (1994). However, the environment dynamics are also assumed to be known in Matinfar and Hashtrudi-Zaad (2005). Recalling LQR in Kirk (1970), it is difficult to find the solution of the Riccati equation if the linear system under study is unknown. Therefore, when the system dynamics are unknown, the methods proposed in Johansson and Spong (1994), Matinfar and Hashtrudi-Zaad (2005) are not applicable. To solve the optimal control problems in the case of unknown system dynamics, adaptive dynamic programming (ADP) or reinforcement learning (RL) has been widely studied in the literature Werbos (1990a,b, 1992), Bertsekas (1995), Werbos (2009), Lewis and Vrabie (2009), Wang et al. (2009). ADP is constructed based on the idea of how biological system interacts with the surrounding environment. In the scheme of ADP, the control system is defined as an actor or agent, modifying its action based on the feedback information of the environment. The actor or agent is rewarded or punished for a control action which is evaluated by a critic Lewis and Vrabie (2009), Wang et al. (2009). Among all ADP approaches, most recognized discrete ADP algorithms are the heuristic dynamic programming (HDP), action-dependent heuristic dynamic programming (ADHDP) or Q-learning Watkins (1989), globalized DHP (GDHP) and dual-heuristic programming (DHP). The common feature of these ADP algorithms is that the design of optimal controller only requires partial information of the system model to be controlled. There are existing works where ADP is adopted for the impedance adaptation of robot arm control. In Kim et al. (2010), natural actor-critic algorithm is adopted and the damping and stiffness matrices are updated according to defined reward functions. In Buchli et al. (2011), the policy improvement with path integrals (PI^2) algorithm is integrated with the reinforcement learning algorithm to achieve variable impedance control. However, as in Arimoto et al. (1984a,b), Cohen and Flash (1991), Yang and

Asada (1996), Tsuji and Morasso (1996), a learning process is still required in Kim et al. (2010), Buchli et al. (2011) for the robot to repeat operations to learn the desired impedance parameters. To solve this problem, this paper aims to develop impedance adaptation in the case of unknown environments dynamics. The method to be developed is based on the latest result of ADP in Jiang and Jiang (2012), where the solution of adaptive optimal control is obtained subject to unknown system dynamics. Two general models of environments are considered, one of which includes damping and stiffness, and the other one includes mass, damping and stiffness. These two models are described as linear systems with unknown dynamics. While ADP in Jiang and Jiang (2012) is only for the state regulation, it is further modified to handle the trajectory tracking. The developed impedance adaptation will result in the desired impedance parameters that are able to guarantee the optimal interaction, subject to unknown environments.

Based on the above discussion, we highlight the contributions of this paper as follows:

- (i) Environment dynamics are taken into consideration in the analysis of optimal robot-environment interaction, and they are described as linear systems with unknown dynamics.
- (ii) ADP for systems with unknown dynamics is modified such that trajectory tracking is achievable and the desired impedance model can be obtained.
- (iii) Impedance parameters of robots are obtained subject to unknown environments, which guarantee the optimal robot-environment interaction in the sense of trajectory tracking and force regulation.

The rest of the paper is organized as follows. In Section 2, the dynamics of the robot and environment are described, and impedance control and the objective of this paper are discussed. In Section 3, impedance adaptation is developed for two general kinds of environments, such that the optimal interaction is achieved subject to unknown environments. In Section 4, the validity of the proposed method is verified through simulation studies. Section 5 concludes this paper.

2 Problem Formulation

2.1 System Description

The system under study includes a rigid robot arm and an environment, where the end-effector of the robot arm physically interacts with the environment. There is a force sensor at the end-effector of the robot arm which measures the interaction force between the robot arm and the environment, as shown in Fig. 1.

Consider the robot kinematics as below

$$x(t) = \phi(q) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ and $q \in \mathbb{R}^n$ are positions/oritations in the Cartesian space and joint coordinates in the joint space, respectively. Differentiating (1) with respect to time results in

$$\dot{x}(t) = J(q)\dot{q} \quad (2)$$

where $J(q) \in \mathbb{R}^{n \times n}$ is the Jacobian matrix.

The robot dynamics in the joint space are given by

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau + J^T(q)f \quad (3)$$

where $M(q) \in \mathbb{R}^{n \times n}$ is the inertia matrix; $C(q, \dot{q})\dot{q} \in \mathbb{R}^n$ denotes the Coriolis and Centrifugal force; $G(q) \in \mathbb{R}^n$ is the gravitational force; $\tau \in \mathbb{R}^n$ is the vector of the control input; and $f(t) \in \mathbb{R}^n$ denotes the force exerted by the environment.

The other part of the system under study is the environment. Without loss of generality, two kinds of environments are considered in this paper, of which the dynamics are respectively described by the

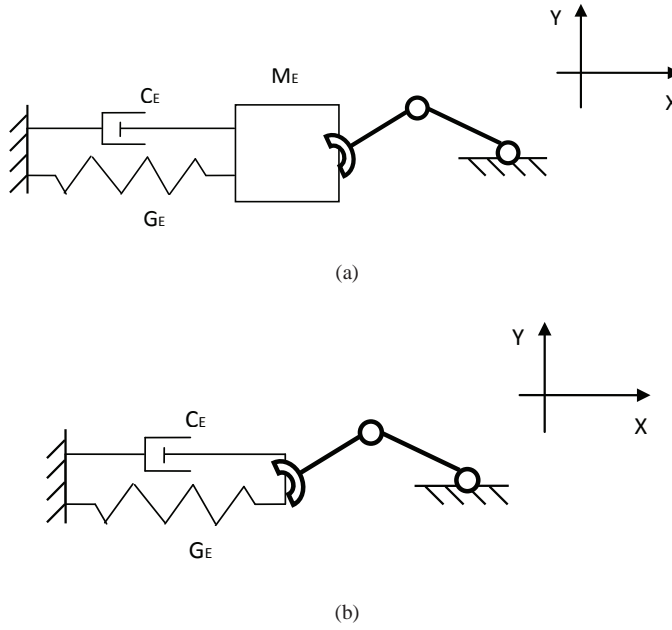


Figure 1.: System under study: (a) mass-damping-stiffness environment, and (b) damping-stiffness environment

following models

$$M_E \ddot{x} + C_E \dot{x} + G_E x = -f \quad (4)$$

$$C_E \dot{x} + G_E x = -f \quad (5)$$

where M_E , C_E and G_E are unknown mass, damping and stiffness matrices of the environment models, respectively. Compared to the second model (5), there is a mass matrix in the first model (4). It will be shown that different impedance models of the robot arm are required for different environments (4) and (5) to achieve the optimal interaction.

Remark 1: The above two kinds of environments represent a large range of environments. For examples, model (4) may describe the dynamics of human limb in physical human-robot interaction Rahman et al. (2002), and model (5) may represent the viscoelastic object in robotic manipulation.

Remark 2: M_E , C_E and G_E are assumed to be unknown constant matrices in this paper. While it is valid in many applications, these matrices can be time-varying in some other applications. The latter assumption makes the problem studied in this paper more complicated and it will be further investigated in the future work.

2.2 Impedance Control

As discussed in the Introduction, impedance control is employed for robots interacting with environments. To implement impedance control, a two-loop control framework is usually adopted, as shown in Fig. 2.

In this framework, the outer-loop is dedicated to generate the virtual desired trajectory in the joint space, i.e., q_d . In particular, the desired impedance model in the Cartesian space is given by

$$f = Z(x_d, x_0) \quad (6)$$

where x_d is the desired trajectory, x_0 is the virtual desired trajectory in the Cartesian space, and $Z(\cdot)$ is a target impedance function to be determined. Then, the virtual desired trajectory in the joint space

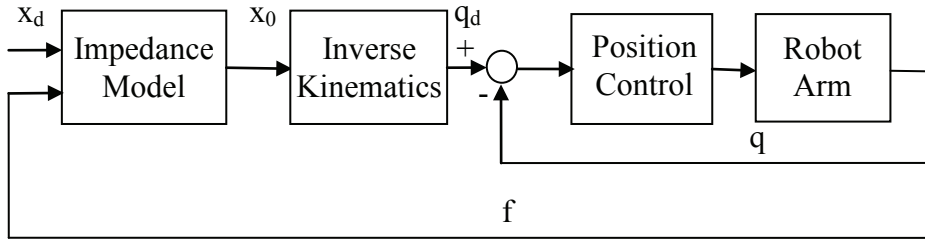


Figure 2.: Impedance Control Diagram

$q_d = \int_0^t J^{-1}(q(v))\dot{x}_0(v)dv$ is obtained according to the interaction force f and the impedance model (6).

Remark 3: Model (6) is a general impedance model which defines the impedance relationship between interaction force and position. In practical implementations, a typical impedance model is usually given by $M_d\ddot{x}_0 + C_d\dot{x}_0 + G_d(x_0 - x_d) = -f$, where M_d , C_d and G_d are desired inertial, damping and stiffness matrices, respectively. A more simplified model, e.g., stiffness model $G_d(x_0 - x_d) = -f$, may be adopted in a specific situation.

The inner-loop is to guarantee the trajectory tracking, i.e., $\lim_{t \rightarrow \infty} q(t) = q_d(t)$. Trajectory tracking of a robot arm has been extensively studied in the literature Ge et al. (1998), and will not be discussed in this paper. For the analysis simplicity, it is assumed that there is an ideal inner-loop position controller such that $q(t) = q_d(t)$ and thus $x(t) = x_0(t)$. In this way, the desired impedance model becomes

$$f = Z(x_d, x) \quad (7)$$

In (7), the impedance function $Z(\cdot)$ is determined such that a certain interaction requirement is satisfied. This is non-trivial considering that the environment dynamics (4) and (5) are unknown. As discussed in the Introduction, iterative learning has been studied to cope with this problem, which however requires repetitive motions. In this paper, we aim to achieve the same objective while avoiding the learning process. This is the motivation to develop impedance adaptation in the rest of this paper.

2.3 Preliminary: Adaptive Optimal Control

The adaptive optimal control proposed in Jiang and Jiang (2012) is briefly introduced in this subsection, of which the results will be used for the development of the impedance adaptation.

Consider the following linear system

$$\dot{\xi} = A\xi + Bu \quad (8)$$

where $\xi \in \mathbb{R}^p$ is the system state, $u \in \mathbb{R}^r$ is the system input, and $A \in \mathbb{R}^{p \times p}$ and $B \in \mathbb{R}^{p \times r}$ are unknown constant matrices.

The following system input

$$u = -K_k \xi \quad (9)$$

minimizes a defined cost function

$$\Gamma = \int_0^\infty [\xi^T Q \xi + u^T R u] dt \quad (10)$$

where $Q \in \mathbb{R}^{p \times p}$ and $R \in \mathbb{R}^{r \times r}$ are the weights of the state and the input which satisfy $Q = Q^T \geq 0$ and $R = R^T > 0$, and K_k with the iteration number k is a matrix obtained by following the procedures as below:

Step 1: Employ $u = K_0\xi + \nu$ as the input on the time interval $[t_0, t_l]$, where K_0 stabilizes the system (8) and ν is the exploration noise to satisfy the persistent excitation (PE) condition. Compute δ_ξ , I_ξ and I_u until the following rank condition is satisfied

$$\text{rank}([I_\xi, I_u]) = \frac{p(p+1)}{2} + pr \quad (11)$$

Step 2: Solve P_k and K_{k+1} according to

$$\begin{bmatrix} \hat{P}_k \\ \text{vec}(K_{k+1}) \end{bmatrix} = (\Theta_k^T \Theta_k)^{-1} \Theta_k^T \Xi_k \quad (12)$$

Step 3: Let $k+1 \rightarrow k$ and repeat Step 2 until $P_k - P_{k-1} \leq \epsilon$ where $\epsilon > 0$ is a predefined constant. And K_k in (9) is obtained.

In Step 1, the following definitions are needed:

$$\begin{aligned} \bar{\xi} &= [\xi_1^2, \xi_1\xi_2, \dots, \xi_1\xi_p, \xi_2^2, \xi_2\xi_3, \dots, \xi_{p-1}\xi_p, \xi_p^2]^T \\ \delta_\xi &= [\bar{\xi}(t_1) - \bar{\xi}(t_0), \bar{\xi}(t_2) - \bar{\xi}(t_1), \dots, \bar{\xi}(t_l) - \bar{\xi}(t_{l-1})]^T \\ I_\xi &= \left[\int_{t_0}^{t_1} \xi \otimes \xi dt, \int_{t_1}^{t_2} \xi \otimes \xi dt, \dots, \int_{t_{l-1}}^{t_l} \xi \otimes \xi dt \right]^T \\ I_u &= \left[\int_{t_0}^{t_1} \xi \otimes u dt, \int_{t_1}^{t_2} \xi \otimes u dt, \dots, \int_{t_{l-1}}^{t_l} \xi \otimes u dt \right]^T \end{aligned} \quad (13)$$

where ξ_i , $i = 1, \dots, p$ are the elements of ξ , l is a positive integer, and “ \otimes ” is the Kronecker product.

In Step 2, the following definitions are needed:

$$\begin{aligned} \hat{P}_k &= [P_{1,1}, 2P_{1,2}, \dots, 2P_{1,p}, P_{2,2}, 2P_{2,3}, \dots, 2P_{p-1,p}, P_{p,p}]^T \\ \Theta_k &= [\delta_\xi, -2I_\xi(I_p \otimes K_k^T R) - 2I_u(I_p \otimes R)] \\ \Xi_k &= -I_\xi \text{vec}(Q_k) \end{aligned} \quad (14)$$

where $P_{i,j}$, $i = 1, \dots, p$, $j = 1, \dots, p$ are the elements of P_k , I_p is a p -dimensional unit matrix, “vec” is the operator that “stretches” a matrix to a vector, and $Q_k = Q + K_k^T R K_k$.

Remark 4: Note that A and B in (8) are not used in the above procedures so the resulted adaptive optimal control is applicable to the system with unknown dynamics. This is a favorable property compared to the traditional LQR since in many situations the system dynamics are difficult to obtain, if not impossible.

3 Impedance Adaptation

This section is dedicated to develop impedance adaptation to achieve optimal interaction subject to unknown environments. Two kinds of environments described by (4) and (5) will be considered in two separate subsections, respectively.

3.1 Mass-Damping-Stiffness Environment

The mass-damping-stiffness environment described by (4) is considered in this subsection. By taking the environment dynamics into consideration, we will develop impedance adaptation such that the

following cost function is minimized

$$\Gamma = \int_0^\infty [\dot{x}^T Q_1 \dot{x} + (x - x_d)^T Q_2 (x - x_d) + f^T R f] dt \quad (15)$$

where $Q_1 \in \mathbb{R}^{n \times n}$ and $Q_2 \in \mathbb{R}^{n \times n}$ are the weights of the velocity and the trajectory tracking error, respectively, and $R \in \mathbb{R}^{n \times n}$ has been defined in (8) and is the weight of the interaction force. Besides, $Q_1 = Q_1^T \geq 0$ and $Q_2 = Q_2^T \geq 0$.

Remark 5: Cost functions similar to (15) have been discussed in the related works Johansson and Spong (1994), Matinfar and Hashtrudi-Zaad (2005), which represent the compromise/combination of the force regulation and trajectory tracking and determine the interaction performance. Note that these cost functions are different from that in the traditional LQR problem, where the cost function usually includes the control input and trajectory tracking error. In the traditional LQR problem, the system under study is the robot itself, while the interaction system under study in this paper includes both the robot and the environment.

Comparing the cost function for a general linear system (10) and the defined cost function in this paper (15), some manipulations are needed to make them identical. In particular, we consider

$$\xi = [\dot{x}^T, x^T, z^T]^T \quad (16)$$

where $z \in \mathbb{R}^m$ is the state of the following system

$$\begin{cases} \dot{z} = Uz \\ x_d = Vz \end{cases} \quad (17)$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times m}$ are two known matrices.

Remark 6: The linear system (17) is to determine the desired trajectory x_d and provides the feasibility to employ the optimal control in trajectory tracking problem. (17) is able to generate a large variety of desired trajectories, e.g., polynomial functions of time of any order.

Then, according to (15) and (17), we have

$$\begin{aligned} \Gamma &= \int_0^\infty (\dot{x}^T Q_1 \dot{x} + [x^T \ x_d^T] \begin{bmatrix} Q_2 & -Q_2 \\ -Q_2 & Q_2 \end{bmatrix} \begin{bmatrix} x \\ x_d \end{bmatrix} + f^T R f) dt \\ &= \int_0^\infty (\dot{x}^T Q_1 \dot{x} + [x^T \ z^T] \begin{bmatrix} Q_2 & -Q_2 V \\ -V^T Q_2 & V^T Q_2 V \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + f^T R f) dt \\ &= \int_0^\infty (\xi^T Q \xi + f^T R f) dt \end{aligned} \quad (18)$$

where $Q = \begin{bmatrix} Q_1 & 0 \\ 0 & Q'_2 \end{bmatrix}$ with $Q'_2 = \begin{bmatrix} Q_2 & -Q_2 V \\ -V^T Q_2 & V^T Q_2 V \end{bmatrix}$.

Considering the defined state (16), we rewrite (4) in the following state-space form

$$\dot{\xi} = A\xi + Bf \quad (19)$$

where $A = \begin{bmatrix} -M_E^{-1} C_E & -M_E^{-1} G_E & 0 \\ I_n & 0 & 0 \\ 0 & 0 & U \end{bmatrix}$ and $B = \begin{bmatrix} -M_E^{-1} \\ 0 \\ 0 \end{bmatrix}$. Note that A and B include the environment dynamics and they are unknown. This is the main reason to cause the difficulty of determining the desired impedance function $Z(\cdot)$ in (7).

If we take the interaction force f in (19) as the “system input” to the environment dynamics, it can be obtained as follows such that the cost function (15) is minimized

$$f = -K_k \xi \quad (20)$$

where K_k is obtained according to the procedures described in Section 2.3.

Remark 7: For the augmented system, both x and \dot{x} are controllable while the uncontrollable state v is asymptotically stable, so the system is stabilizable.

To understand (20) in the sense of impedance control, we assume that the optimal control has been obtained, which is

$$f = -K\xi = -R^{-1}B^T P\xi \quad (21)$$

where K is the optimal feedback gain matrix and $P = P^T \in \mathbb{R}^{(2n+m) \times (2n+m)}$ is the solution of the following Riccati equation

$$PA + A^T P - PBR^{-1}B^T P + Q = 0 \quad (22)$$

Denote

$$P = \begin{bmatrix} P_1 & P_2 & P_3 \\ * & * & * \\ * & * & * \end{bmatrix} \quad (23)$$

where $P_1 \in \mathbb{R}^{n \times n}$, $P_2 \in \mathbb{R}^{n \times n}$, $P_3 \in \mathbb{R}^{n \times m}$ and “*” stand for elements which we do not care about. Substituting (23) into (21) leads to

$$\begin{aligned} f &= -R^{-1}P_1\dot{x} - R^{-1}P_2x - R^{-1}P_3z \\ &= -R^{-1}P_1\dot{x} - R^{-1}P_2x - R^{-1}P_4x_d \end{aligned} \quad (24)$$

where $P_4 = P_3(V^T V)^{-1}V^T$.

Comparing the above equation (24) with the desired impedance model (7), the exact impedance function which guarantees the optimal interaction is obtained. Recalling the implementation of impedance control as described in Section 2.2, the virtual desired trajectory in the Cartesian space x is obtained according to (24) with measured f and given x_d , and the inner-position control loop is to guarantee the trajectory tracking in the joint space. In this way, the optimal interaction is achieved and (24) is the resulted impedance function in the presence of unknown environment dynamics.

3.2 Damping-Stiffness Environment

For the environment without mass, i.e., the damping-stiffness environment described by (5), the resulted impedance adaptation is different and it is discussed in this subsection.

First, we consider to minimize the following cost function

$$\Gamma' = \int_0^\infty [(x - x_d)^T Q_2 (x - x_d) + f^T R f] dt \quad (25)$$

Remark 8: Note that the component to penalize the velocity in (15) has disappeared in (25). The reason is that the velocity is not the state of the environment (5), which will be further explained in the following.

Correspondingly, we consider the state $\xi' = [x^T, z^T]^T$, where $z \in \mathbb{R}^m$ has the same meaning as defined in (17). Based on the similar manipulation as in the previous section, (25) can be rewritten as

$$\Gamma' = \int_0^\infty (\xi'^T Q \xi' + f^T R f) dt \quad (26)$$

where $Q' = \begin{bmatrix} Q_2 & -Q_2 V \\ -V^T Q_2 & V^T Q_2 V \end{bmatrix}$.

Similarly, (5) can be rewritten in the following state-space form

$$\dot{\xi}' = A' \xi' + B' f \quad (27)$$

where $A' = \begin{bmatrix} -C_E^{-1} G_E & 0 \\ 0 & U \end{bmatrix}$ and $B' = \begin{bmatrix} -C_E^{-1} \\ 0 \end{bmatrix}$.

If we take the interaction force f in (27) as the “system input” to the environment dynamics (5), it can be obtained as follows such that the cost function (25) is minimized

$$f = -K'_k \xi' \quad (28)$$

where K'_k is obtained according to the procedures described in Section 2.3.

To understand (28) in the sense of impedance control, we assume that the optimal control has been obtained, which is

$$f = -K' \xi' = -R^{-1} B'^T P' \xi' \quad (29)$$

where K' is the optimal feedback gain matrix and $P' = P'^T \in \mathbb{R}^{(n+m) \times (n+m)}$ is the solution of the following Riccati equation

$$P' A' + A'^T P' - P' B' R^{-1} B'^T P' + Q' = 0 \quad (30)$$

Denote

$$P' = \begin{bmatrix} P'_1 & P'_2 \\ * & * \end{bmatrix} \quad (31)$$

where $P'_1 \in \mathbb{R}^{n \times n}$ and $P'_2 \in \mathbb{R}^{n \times m}$. Substituting (31) into (29) leads to

$$\begin{aligned} f &= -R^{-1} P'_1 x - R^{-1} P'_2 z \\ &= -R^{-1} P'_1 x - R^{-1} P'_3 x_d \end{aligned} \quad (32)$$

where $P'_3 = P'_2 (V^T V)^{-1} V^T$.

It is found that the resulted control is variable stiffness control, i.e., there is no damping component as in (24). In this sense, we conclude that both the damping and stiffness components are needed in the impedance adaptation for the optimal interaction with the mass-damping-stiffness environments, and only the stiffness component is needed for the optimal interaction with the damping-stiffness environments. Similarly as in Section 3.1, the desired impedance function $Z(\cdot)$ in (7) is obtained as (32), which guarantees the optimal interaction subject to unknown environment dynamics (5).

Remark 9: The adaptive optimal control proposed in Jiang and Jiang (2012) is limited to linear and time invariant LQR problem, so the methods developed in Sections 3.1 and 3.2 are also limited to linear and time invariant environment dynamics. However, these methods could also be further modified to suit the case where the parameters of the environment dynamics are changing slowly, using a time varying parameter estimation.

4 Simulation Study

4.1 Simulation Conditions

In this section, we consider a robot arm with two revolute joints physically interacting with two environments, as discussed through this paper and shown in Figs. 1(a) and 1(b). The simulation is conducted with the Robotics Toolbox Corke (1996).

The parameters of the robot arm are: $m_1 = m_2 = 2.0\text{kg}$, $l_1 = l_2 = 0.2\text{m}$, $i_1 = i_2 = 0.027\text{kgm}^2$, $l_{c1} = l_{c2} = 0.1\text{m}$, where m_j, l_j, i_j, l_{cj} , $j = 1, 2$, represent the mass, the length, the inertia about the z-axis that comes out of the page passing through the center of mass, and the distance from the previous joint to the center of mass of link i , respectively.

The initial joint coordinates of the robot arm are $q_1 = -\frac{\pi}{3}$ and $q_2 = \frac{2\pi}{3}$, and thus the initial position in the Cartesian space is $x_d = [0.2 \ 0]^T$. The desired trajectory in the Cartesian space is determined by (17) with $U = 1$ and $V = 0.3$. It is assumed that the force exerted by the environment is only in X direction and thus the robot arm in Y direction is interaction-free. Nevertheless, the inner-loop position control is designed for both joints. In particular, adaptive control in Slotine and Li (1987) is adopted for the inner-loop position control.

According to the adaptation procedure in Section 3, three steps are included in a single simulation. As discussed in Section 2.3, the exploration noise which meets the persistent excitation condition is required to guarantee the estimated parameters to converge to the true values. It is well known that the sum of sinusoid waves is a typical swept frequency signal which meets the persistent excitation condition. Therefore, in the simulation, the exploration noise is chosen as $\nu = \sum_{w=1}^{10} \frac{0.4}{w} \sin(wt)$ and added in the first step. The impedance model for this step is $f = f_0 + \nu$, where $f_0 = -\dot{x} - x + x_d$ in Section 4.2 and $f_0 = -x + x_d$ in Section 4.3. Impedance adaptation is conducted in the second step, and it stops when the condition $\|P_k\| < 10^{-10}$ satisfies. The initial P_k is $P_0 = 10I_p$, where I_p represents the p dimensional unit matrix. The impedance function for this step is f_0 . The desired impedance model based on the proposed method is obtained at the end of the second step, and it is used in the third step.

Corresponding to Sections 3.1 and 3.2, the following two environments are considered in two subsections: $0.001\ddot{x} + 0.01\dot{x} + (x - 0.2) = -f$ and $0.01\ddot{x} + (x - 0.2) = -f$. Note that 0.2 is the initial position of the robot arm. According to (22) and (30), if A and B are known, the optimal solutions of the Riccati equations can be obtained and thus the desired impedance model (24) and (32) can be obtained. It is referred as ‘‘LQR’’, and compared with the the proposed method in this paper. It is necessary to emphasize that A , B and thus the desired impedance model are only available in the simulation studies for the comparison purpose, and they are usually unknown or need to be estimated in a typical application. This is the motivation of this paper, as already discussed in the Introduction.

4.2 Mass-Damping-Stiffness Environment

In this subsection, the mass-damping-stiffness environment $0.001\ddot{x} + 0.01\dot{x} + (x - 0.2) = -f$ is considered. In the first case, the weights in (15) are given by $Q_1 = 1$, $Q_2 = 1$ and $R = 1$. The desired impedance model (24) is $f = -0.99\dot{x} - 0.41x + 0.04x_d$ based on known A and B . The simulation results are shown in Figs. 3, 4, 5 and 6. In Fig. 3, the position of the robot arm in the Cartesian space is shown. In the first two seconds, there is a large position error between LQR and the proposed method. This is because there is exploration noise and the initial impedance model $f = -\dot{x} - x + x_d + \nu$ is not the desired model. The second step takes a very short time and the desired impedance function converges very quickly. More details about the convergence performance can be found in Fig. 5, where the error of impedance parameters with respect to iteration number is shown. This error is defined as $\|K_k - K\|$, and it decreases to around 0.05 when the adaptation stops at the 5th iteration (each iteration takes 0.1s). The desired impedance model is obtained as $f = -0.98\dot{x} - 0.38x + 0.06x_d$ with the proposed method, and it is used until the end of the simulation. It is found that the obtained impedance model is very near to but not exactly the same as the desired one under LQR, i.e., $f = -0.99\dot{x} - 0.41x + 0.04x_d$. As a result, the position in Fig. 3 is near to the position under LQR but there is still a small error. This is caused by the adaptation process in the inner-loop, and it is illustrated by Fig. 6 where the control performance of the

inner-loop is shown. Particularly, although zero tracking error in the inner-loop is achieved after about 2s, the tracking error exists at the beginning of the simulation. Therefore, the dynamics of the robot arm are actually not exactly governed by the given impedance model.

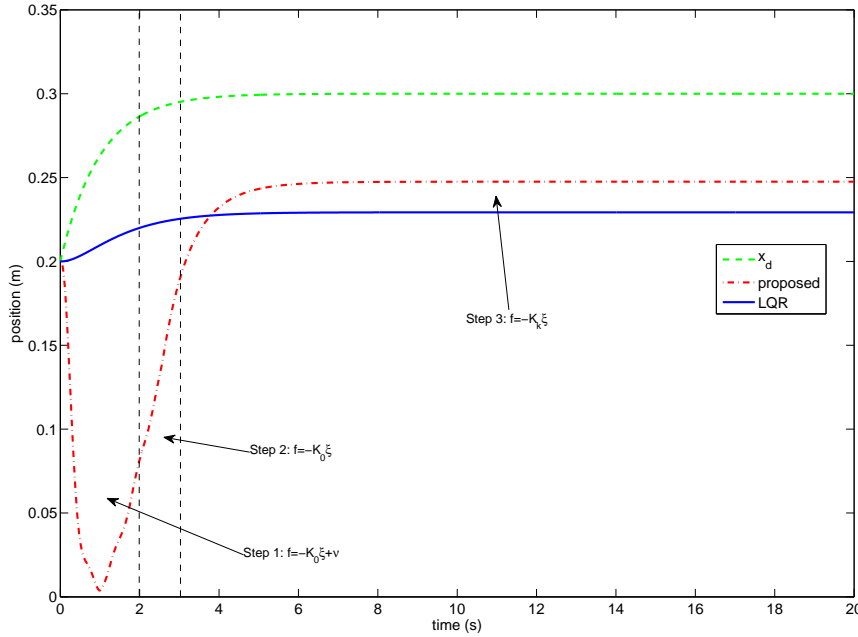


Figure 3.: Desired trajectory and actual trajectory, $Q_1 = 1$, $Q_2 = 1$ and $R = 1$

To further investigate the effect of inner-loop performance on the impedance adaptation, an additional simulation study is conducted. Instead of starting from the beginning of the simulation, the impedance adaptation starts at 2s when the inner-loop tracking error is near to zero. Simulation results are shown in Figs. 7 and 8. The desired impedance model is obtained as $f = -0.98\dot{x} - 0.40x + 0.04x_d$ with the proposed method, which shows that the perfect adaptation can be achieved when a perfect tracking is guaranteed throughout the parameter adaptation period. By comparing the simulation results in Figs. 3, 4 and Figs. 7, 8, we conclude that a good inner-loop performance during the parameter adaptation period is crucial to guarantee a perfect impedance adaptation. When the good control performance of the inner-loop is not achieved during the parameter adaptation period, the adapted parameters cannot converge to the exact values. Conversely, when a perfect inner-loop control performance is guaranteed during the adaptation process, the perfect impedance adaptation can be realized.

To further verify the effectiveness of the proposed impedance adaptation, another cost function is chosen in the second case. Particularly, the weights in (15) are given by $Q_1 = 1$, $Q_2 = 10$ and $R = 1$. Compared to that in the first case, the weight of the velocity is larger so it is expected that the system response is slower. Similarly, the desired impedance model (24) is obtained as $f = -3.15\dot{x} - 0.41x + 0.02x_d$ based on known A and B . The simulation results in this case are given in Figs. 9 and 10, and the impedance model obtained with the proposed method is $f = -3.14\dot{x} - 0.40x + 0.02x_d$. While the position in Fig. 3 converges to around 0.25m in 6s, the position in Fig. 9 converges to around 0.22m in 10s. Similarly, the interaction force in Fig. 4 converges to around -0.05N in 8s, and in Fig. 10 it converges to around -0.02N in 16s. The above results are coherent with the expectation and verify the validity of the propose method. Similarly as in LQR, different Q_1 , Q_2 and R can be chosen to realize different interaction performances, e.g., either “softer” interaction or more accurate trajectory tracking.

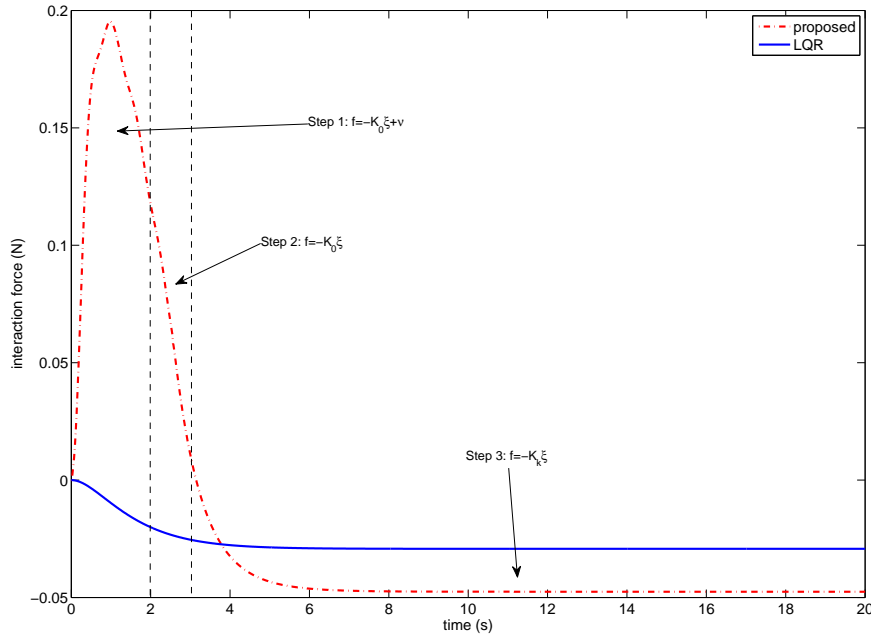


Figure 4.: Interaction force, $Q_1 = 1$, $Q_2 = 1$ and $R = 1$

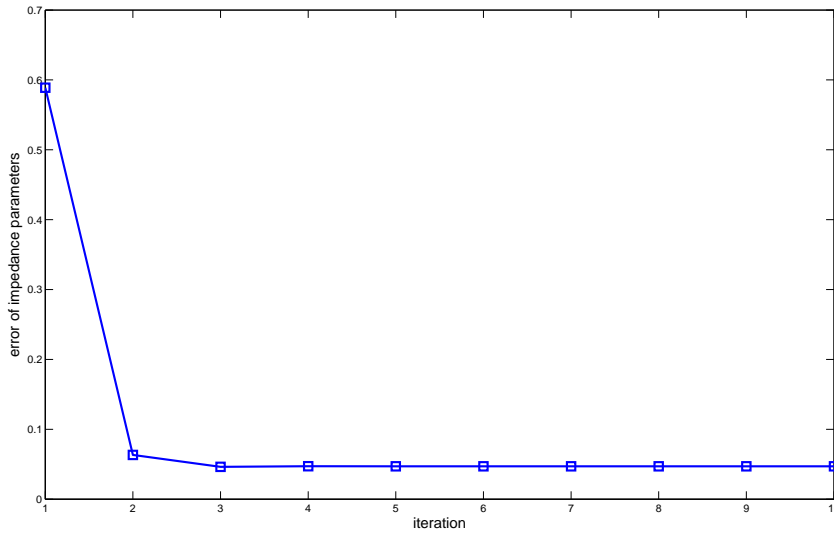


Figure 5.: Error of impedance parameters, $Q_1 = 1$, $Q_2 = 1$ and $R = 1$

4.3 Damping-Stiffness Environment

The damping-stiffness environment $0.01\dot{x} + (x - 0.2) = -f$ is considered in this subsection. The simulation conditions are the same as in the previous subsection, which are described in Section 4.1. The weights in (25) are given by $Q_2 = 1$ and $R = 1$ and the desired impedance model (32) is obtained as $f = -0.41x + 0.70x_d$ based on known A' and B' . Similar results as in the previous subsection are obtained. In particular, the impedance model obtained with the proposed method is $f = -0.40x + 0.71x_d$. The optimal interaction performance is achieved as shown in Figs. 11 and 12.

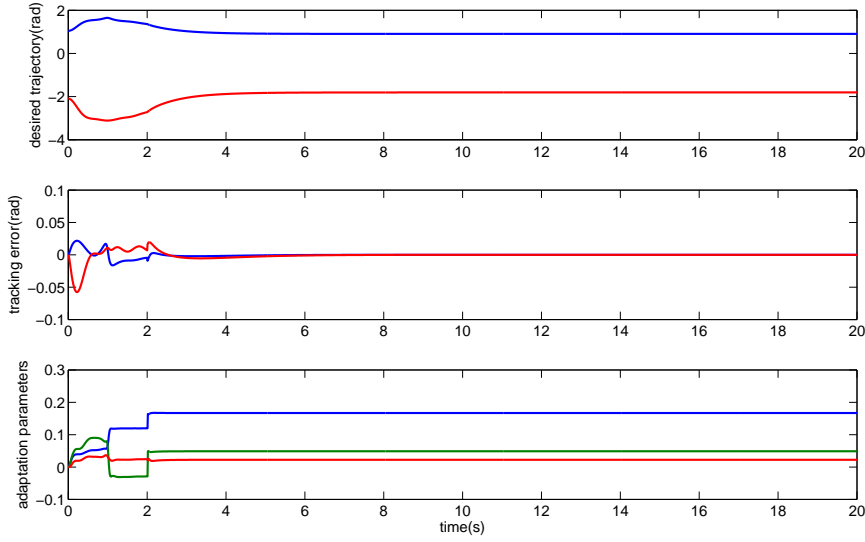
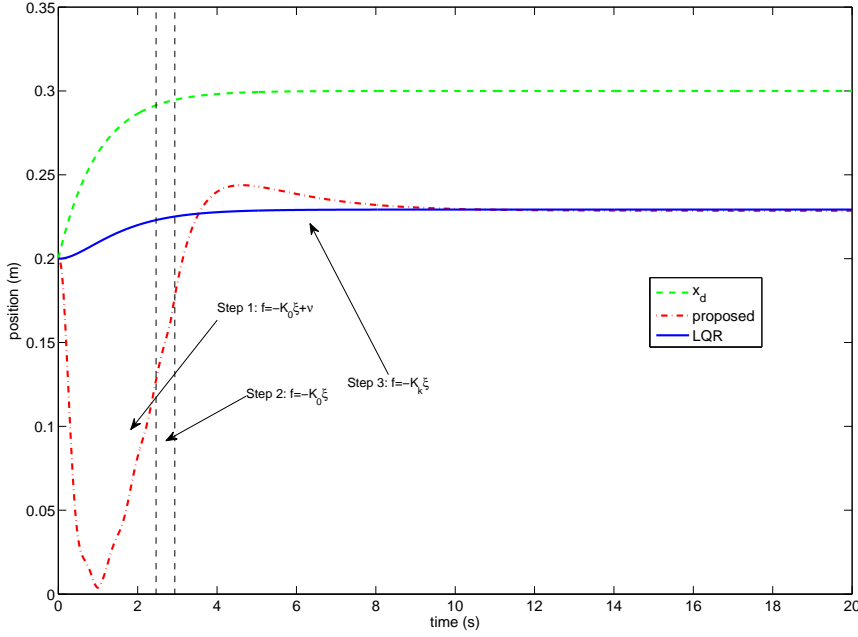


Figure 6.: Inner-loop control performance

Figure 7.: Desired trajectory and actual trajectory, $Q_1 = 1$, $Q_2 = 1$ and $R = 1$

4.4 Discussion

In summary, it has been shown that the proposed method can be adopted to obtain a desired impedance model based on a given cost function which determines an optimal interaction performance. Compared to impedance learning developed in the literature such as Kim et al. (2010), Buchli et al. (2011), the proposed method does not require the repetitive learning process and thus provides certain convenience. The optimal impedance adaptation can be achieved if the “perfect” tracking can be guaranteed in the inner-loop. As discussed in Remark 9, the proposed method in Jiang and Jiang (2012) may not be applicable to the scenario where the environment is changing rapidly. In the future work, we will investigate how to derive an impedance adaptation in face of dynamically changing environments. Besides, some issues

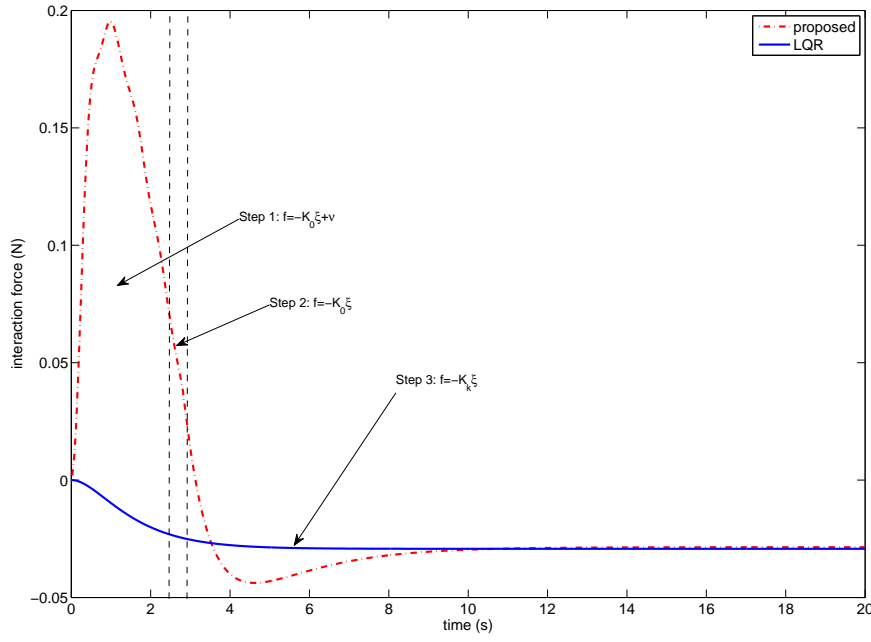


Figure 8.: Interaction force, $Q_1 = 1$, $Q_2 = 1$ and $R = 1$

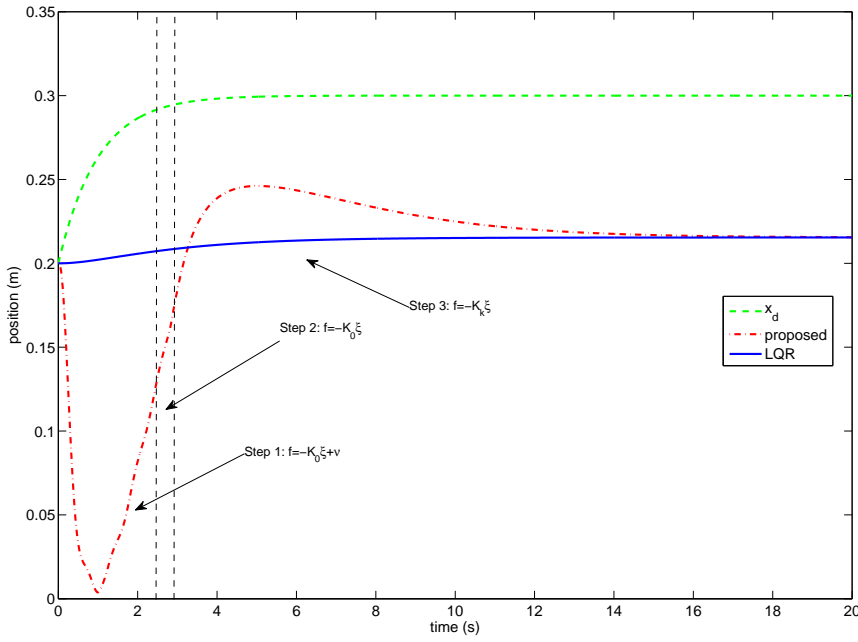


Figure 9.: Desired trajectory and actual trajectory, $Q_1 = 1$, $Q_2 = 10$ and $R = 1$

in real-world implementations might not be well described in the simulation studies, so future work will also be dedicated to incorporate the proposed impedance adaptation with a real robot arm and further investigate the issues in real-world implementations (e.g., the effect of exploration noise, environment model uncertainties and time delay) of the proposed impedance adaptation. Furthermore, as discussed in Braun et al. (2012b), it is nontrivial to find a proper cost function in many situations, which will be also one of the future works that we will focus on.

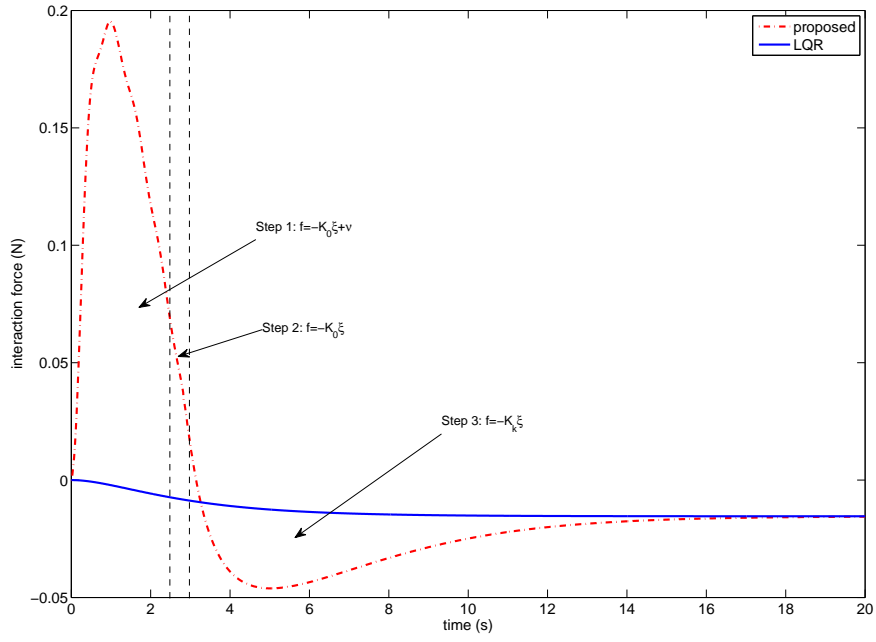


Figure 10.: Interaction force, $Q_1 = 1$, $Q_2 = 10$ and $R = 1$

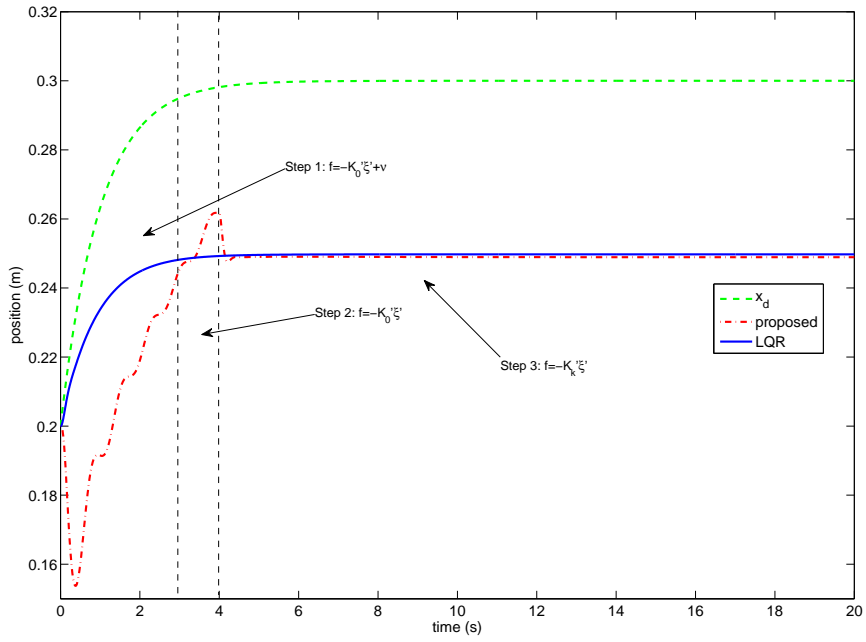


Figure 11.: Desired trajectory and actual trajectory, $Q_2 = 1$ and $R = 1$

5 Conclusion

In this paper, impedance adaptation has been developed to obtain the desired impedance parameters such that the optimal interaction is realized subject to unknown environments. The dynamics of the unknown environments have been investigated and the interaction requirement has been described by minimizing a certain cost function which includes trajectory tracking and force regulation. Adaptive op-

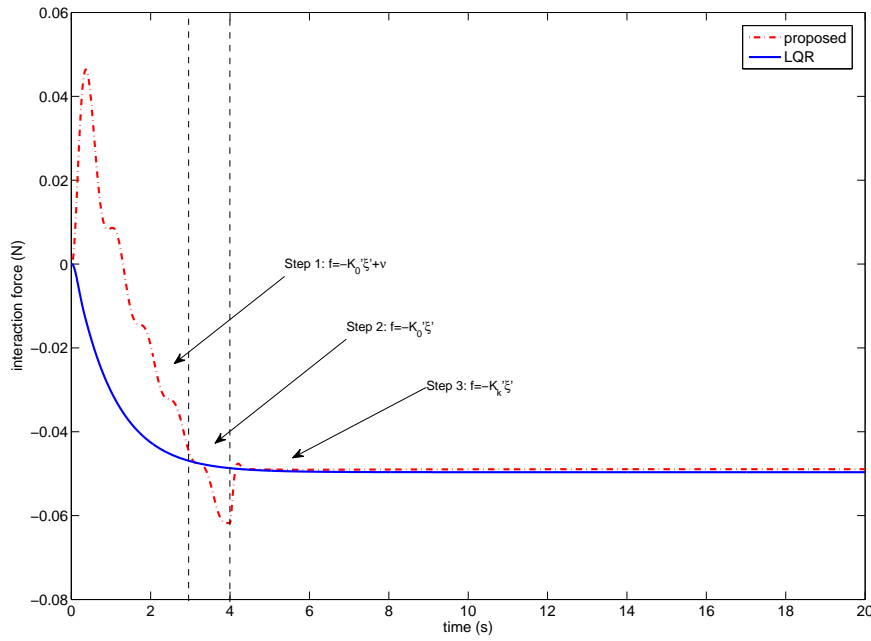


Figure 12.: Interaction force, $Q_2 = 1$ and $R = 1$

timal control for unknown linear system has been employed as the fundamental of the proposed method. The validity of the proposed method has been verified through simulation studies.

References

- Arimoto, S., Kawamura, S., and Miyazaki, F. (1984a), “Bettering operation of dynamic systems by learning: A new control theory for servomechanism or mechatronics systems,” *Proceedings of the IEEE Conference on Decision and Control*, 23, 1064–1069.
- Arimoto, S., Kawamura, S., and Miyazaki, F. (1984b), “Bettering operation of robots by learning,” *Journal of Robotic Systems*, 1, 123–140.
- Bertsekas, D., *Dynamic Programming and Optimal Control*, Vol. 1, Athena Scientific Belmont, MA (1995).
- Braun, D.J., Howard, M., and Vijayakumar, S. (2012b), “Optimal variable stiffness control: formulation and application to explosive movement tasks,” *Autonomous Robots*, 33, 237–253.
- Braun, D.J., Petit, F., Huber, F., Haddadin, S., Smagt, P., Albu-Schaffer, A., and Vijayakumar, S. (2012a), “Optimal Torque and Stiffness Control in Compliantly Actuated Robots,” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2801–2808.
- Buchli, J., Stulp, F., Theodorou, E., and Schaal, S. (2011), “Learning variable impedance control,” *International Journal of Robotics Research*, 30, 820–833.
- Buerger, S.P., and Hogan, N. (2007), “Complementary stability and loop shaping for improved human-robot interaction,” *IEEE Transactions on Robotics*, 23, 232–244.
- Cheah, C.C., and Wang, D. (1998), “Learning impedance control for robotic manipulators,” *IEEE Transactions on Robotics and Automation*, 14, 452–465.
- Cohen, M., and Flash, T. (June 1991), “Learning impedance parameters for robot control using an associative search network,” *IEEE Transactions on Robotics and Automation*, 7, 382–390.
- Colbaugh, R., Seraji, H., and Glass, K. (1991), “Direct adaptive impedance control of manipulators,” *Proceedings of the IEEE Conference on Decision and Control*, pp. 2410–2415.
- Colgate, J.E., and Hogan, N. (1988), “Robust control of dynamically interacting systems,” *International Journal of Control*, 48, 65–88.
- Corke, P.I. (1996), “A Robotics Toolbox for MATLAB,” *IEEE Robotics and Automation Magazine*, 3, 24–32.
- Craig, J.J., and Raibert, M.H. (1979), “A Systematic Method of Hybrid Position/Force Control of a Manipulator,” *Computer Software and Applications Conference, IEEE Computer Society*, pp. 446–451.
- De Roover, D., Bosgra, O., and Steinbuch, M. (2000), “Internal-model-based design of repetitive and iterative

- learning controllers for linear multivariable systems,” *International Journal of Control*, 73, 914–929.
- Ge, S.S., Lee, T.H., and Harris, C.J., *Adaptive Neural Network Control of Robotic Manipulators*, London: World Scientific (1998).
- Hogan, N. (1985), “Impedance control: an approach to manipulation-Part I: Theory; Part II: Implementation; Part III: Applications,” *Transaction ASME J. Dynamic Systems, Measurement and Control*, 107, 1–24.
- Huang, L., Ge, S.S., and Lee, T.H. (2006), “Position/force control of uncertain constrained flexible joint robots,” *Mechatronics*, 16, 111–120.
- Jiang, Y., and Jiang, Z.P. (2012), “Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics,” *Automatica*, 48, 2699–2704.
- Johansson, R., and Spong, M.W. (1994), “Quadratic Optimization of Impedance Control,” *Proceedings of IEEE International Conference of Robotics and Automation*, 1, 616–621.
- Kim, B., Park, J., Park, S., and Kang, S. (2010), “Impedance learning for robotic contact tasks using natural actor-critic algorithm,” *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, 40, 433–443.
- Kirk, D.E., *Optimal Control Theory: An Introduction*, Prentice-Hall Inc., Englewood Cliffs New Jersey (1970).
- Lewis, F., and Vrabie, D. (2009), “Reinforcement learning and adaptive dynamic programming for feedback control,” *Circuits and Systems Magazine, IEEE*, 9, 32–50.
- Li, Y., Ge, S.S., and Yang, C. (2012), “Learning Impedance Control for Physical Robot-Environment Interaction,” *International Journal of Control*, 85, 182–193.
- Li, Z.J., Ge, S.S., and Ming, A. (2007), “Adaptive robust motion/force control of holonomi-constrained nonholonomic mobile manipulators,” *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, 37, 607–616.
- Lu, W.S., and Meng, Q.H. (1991), “Impedance control with adaptation for robotic manipulations,” *IEEE Transactions on Robotics and Automation*, 7, 408–415.
- Matinfar, M., and Hashtrudi-Zaad, K. (2005), “Optimization-based Robot Compliance Control: Geometric and Linear Quadratic Approaches,” *The International Journal of Robotics Research*, 24, 645–656.
- Rahman, M.M., Ikeura, R., and Mizutani, K. (2002), “Investigation of the impedance characteristic of human arm for development of robots to cooperate with humans,” *JSME International Journal Series C*, 45, 510–518.
- Slotine, J.J.E., and Li, W. (1987), “On the adaptive control of robotic manipulators,” *The International Journal of Robotics Research*, 6.
- Stanisic, R.Z., and Fernandez, A.V. (2012), “Adjusting the parameters of the mechanical impedance for velocity, impact and force control,” *Robotica*, 30, 583–597.
- Tsuji, T., and Morasso, P.G. (1996), “Neural Network learning of robot arm impedance in operational space,” *IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics*, 26, 290–298.
- Tsumugiwa, T., Yokogawa, R., and Hara, K. (2001), “Variable impedance control with regard to working process for man-machine cooperation-work system,” *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1564–1569.
- Tsumugiwa, T., Yokogawa, R., and Hara, K. (2002), “Variable impedance control based on estimation of human arm stiffness for human-robot cooperative calligraphic task,” *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 644–650.
- Uemura, M., and Kawamura, S. (2009), “Resonance-based Motion Control Method for Multi-Joint Robot through Combining Stiffness Adaptation and Iterative Learning Control,” *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1543–1548.
- Wang, D., and Cheah, C.C. (1998), “An iterative learning-control scheme for impedance control of robotic manipulators,” *The International Journal of Robotics Research*, 17, 1091–1099.
- Wang, F., Zhang, H., and Liu, D. (2009), “Adaptive dynamic programming: An introduction,” *Computational Intelligence Magazine*, 4, 39–47.
- Watkins, C. (1989), “Learning from delayed rewards,” *Cambridge University, Cambridge, England, Doctoral thesis*.
- Werbos, P.J. (1990b), “Consistency of HDP applied to a simple reinforcement learning problem,” *Neural Networks*, 3, 179–189.
- Werbos, P.J. (1990a), “A menu of designs for reinforcement learning over time,” *Neural Networks for Control*, pp. 67–95.
- Werbos, P.J., *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, Vol. 15 (1992).
- Werbos, P.J. (2009), “Intelligence in the brain: A theory of how it works and how to build it,” *Neural Networks*, 22, 200–212.
- Xu, J., Viswanathan, B., and Qu, Z. (2000), “Robust learning control for robotic manipulators with an extension to a class of non-linear systems,” *International Journal of Control*, 73, 858–870.
- Yang, B.H., and Asada, H. (1996), “Progressive learning and its application to robot impedance learning,” *IEEE Transactions on Neural Networks*, pp. 941–952.